Gain Matrix vs Weighted Mean

This note follows the notation of Billoir, NIM **225** (1984) 352. We will consider a fit with N points. This discussion considers the case that the information at points $n, \dots N$ has already been added to the fit and that the information at point $n-1$ is about to be added. First, define the following symbols:

- $\boldsymbol{\eta}$ The optimal estimator of the state vector at the point $n-1$, using all of the information from $n-1, n, \dots, N$. This is what we want to solve for when we add information at the point $n-1$.
- $\boldsymbol{\eta}'_{n-1}$ The estimator of the state vector at $n-1$, using only the information from $n, \dots, N$.
- $\mathbf{V}'_{n-1}$ The covariance matrix of $\boldsymbol{\eta}'_{n-1}$.
- d The optimal estimator for the drift distance at $n-1$, using all of the information from $n-1, n, \dots N$.
- d'_{n-1} The estimator for the drift distance at $n-1$, using only the information from $n, \dots N$.
- d^m_{n-1} The measured drift distance at the point $n-1$.
- σ^m_{n-1} The measurement error on d^m_{n-1}.
- σ'_{n-1} The error on d'_{n-1}.

In order to add information to the fit at point $n-1$, we want to minimize the quantity,

$$(\boldsymbol{\eta} - \boldsymbol{\eta}'_{n-1})^T (\mathbf{V}'_{n-1})^{-1} (\boldsymbol{\eta} - \boldsymbol{\eta}'_{n-1}) + \frac{(d - d^m_{n-1})^2}{(\sigma^m_{n-1})^2}. \tag{1}$$

It will be minimized with respect to $\boldsymbol{\eta}$. Billoir's equation (3) gives the conditions for the minimum when the measurement is one of the elements of the state vector (ie one of the track parameters). When the measurement is not one of the elements of the state vector, the minimum is given by,

$$\left(\mathbf{V}'_{n-1}\right)^{-1} (\boldsymbol{\eta} - \boldsymbol{\eta}'_{n-1}) + \left(\frac{\partial d}{\partial \eta_k}\right) \frac{d - d^m_{n-1}}{(\sigma^m_{n-1})^2} = 0. \tag{2}$$

Now, define the column vector $\mathbf{A}$ as,

$$A_k = \left.\frac{\partial d}{\partial \eta_k}\right|_{d=d'_{n-1}}. \tag{3}$$

Then expand d around d'_{n-1},

$$d = d'_{n-1} + \mathbf{A}^T (\boldsymbol{\eta} - \boldsymbol{\eta}'_{n-1}). \tag{4}$$

With this substitution, equation 2 becomes,

$$\left[\left(\mathbf{V}'_{n-1}\right)^{-1}+\frac{\mathbf{A}\mathbf{A}^T}{(\sigma^m_{n-1})^2}\right](\boldsymbol{\eta}-\boldsymbol{\eta}'_{n-1})=\mathbf{A}\frac{d^m_{n-1}-d'_{n-1}}{(\sigma^m_{n-1})^2}. \tag{5}$$

This is easily solved with the help of the identity,

$$\left(\mathbf{V}^{-1}+\frac{\mathbf{A}\mathbf{A}^T}{\sigma^2}\right)^{-1} = \mathbf{V}-\frac{(\mathbf{V}\mathbf{A})(\mathbf{V}\mathbf{A})^T}{\sigma^2+\mathbf{A}^T\mathbf{V}\mathbf{A}}. \tag{6}$$

We also identify,

$$(\sigma'_{n-1})^2 = \mathbf{A}^T\mathbf{V}'_{n-1}\mathbf{A}. \tag{7}$$

The solution to equation 5 is then,

$$(\boldsymbol{\eta}-\boldsymbol{\eta}'_{n-1}) = \frac{\mathbf{V}'_{n-1}\mathbf{A}}{(\sigma^m_{n-1})^2+(\sigma'_{n-1})^2}(d^m_{n-1}-d'_{n-1}). \tag{8}$$

Now consider the special case that the measurement is one of the elements of the state vector, for example $\mathbf{A}^T=(1,0,0,0,0)$. In this case, $(\sigma'_{n-1})^2=(\mathbf{V}'_{n-1})_{11}$ and equation 8 becomes,

$$(\boldsymbol{\eta}-\boldsymbol{\eta}'_{n-1})_k = \frac{(\mathbf{V}'_{n-1})_{k1}}{(\sigma^m_{n-1})^2+(\sigma'_{n-1})^2}(d^m_{n-1}-d'_{n-1}). \tag{9}$$

In particular,

$$(\boldsymbol{\eta}-\boldsymbol{\eta}'_{n-1})_1 = \frac{(\sigma'_{n-1})^2}{(\sigma^m_{n-1})^2+(\sigma'_{n-1})^2}(d^m_{n-1}-d'_{n-1}). \tag{10}$$

Now consider the formula for the mean of two measurements,

$$\overline{x} = (\frac{x_1}{\sigma_1}+\frac{x_2}{\sigma_2})/(\frac{1}{\sigma_1}+\frac{1}{\sigma_2}) \tag{11}$$

$$\overline{x}-x_1 = \frac{\sigma_1^2}{\sigma_2^2+\sigma_1^2}(x_2-x_1) \tag{12}$$

A comparison of equations 10 and 12 shows that, when the measurement corresponds to one element in the state vector, then the optimal estimator of that element in the state vector has a simple interpretation. It is the weighted mean of two quantities, the present measurement and the prediction from all previous measurements. When the measurement does not correspond to an element of the state vector, then the optimal estimator is given by equation 8. This too can be interpreted as a sort of weighted mean.